

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/991,212

TIME: 10:22:21

Input Set : N:\Crf3\RULE60\09991212.txt

Output Set: N:\CRF3\01152002\I991212.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Lal, Preeti

6 Bandman, Olga

8 (ii) TITLE OF INVENTION: NOVEL HUMAN SODIUM-DEPENDENT

9 PHOSPHATE CO-TRANSPORTER

11 (iii) NUMBER OF SEQUENCES: 7

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

15 (B) STREET: 3174 Porter Drive

16 (C) CITY: Palo Alto

17 (D) STATE: CA

18 (E) COUNTRY: US

19 (F) ZIP: 94304

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Diskette

23 (B) COMPUTER: IBM Compatible

24 (C) OPERATING SYSTEM: DOS

25 (D) SOFTWARE: FastSEQ Version 2.0

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/09/991,212

C--> 29 (B) FILING DATE: 16-Nov-2001

30 (C) CLASSIFICATION:

32 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: 09/391,958

34 (B) FILING DATE:

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Billings, Lucy J.

39 (B) REGISTRATION NUMBER: 36,749

40 (C) REFERENCE/DOCKET NUMBER: PF-0221 US

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 415-855-0555

44 (B) TELEFAX: 415-845-4166

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 401 amino acids

51 (B) TYPE: amino acid

52 (C) STRANDEDNESS: single

53 (D) TOPOLOGY: linear

55 (vii) IMMEDIATE SOURCE:

56 (A) LIBRARY: BRAITUT02

57 (B) CLONE: 754412

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 Met Gln Val Asp Glu Thr Leu Ile Pro Arg Lys Val Pro Ser Leu Cys

62 1 5 10 15

63 Ser Ala Arg Tyr Gly Ile Ala Leu Val Leu His Phe Cys Asn Phe Thr

64 20 25 30

ENTERED

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/991,212

TIME: 10:22:21

Input Set : N:\Crif3\RULE60\09991212.txt

Output Set: N:\CRF3\01152002\I991212.raw

```

65 Thr Ile Ala Gln Asn Val Ile Met Asn Ile Thr Met Val Ala Met Val
66          35          40          45
67 Asn Ser Thr Ser Pro Gln Ser Gln Leu Asn Asp Ser Ser Glu Val Leu
68          50          55          60
69 Pro Val Asp Ser Phe Gly Gly Leu Ser Lys Ala Pro Lys Ser Leu Pro
70          65          70          75          80
71 Ala Lys Ser Ser Ile Leu Gly Gly Gln Phe Ala Ile Trp Glu Arg Trp
72          85          90          95
73 Gly Pro Pro Gln Glu Arg Ser Arg Leu Cys Ser Ile Ala Leu Ser Gly
74          100          105          110
75 Met Leu Leu Gly Cys Phe Thr Ala Ile Leu Ile Gly Gly Phe Ile Ser
76          115          120          125
77 Glu Thr Leu Gly Trp Pro Phe Val Phe Tyr Ile Phe Gly Gly Val Gly
78          130          135          140
79 Cys Val Cys Cys Leu Leu Trp Phe Val Val Ile Tyr Asp Asp Pro Val
80          145          150          155          160
81 Ser Tyr Pro Trp Ile Ser Thr Ser Glu Lys Glu Tyr Ile Ile Ser Ser
82          165          170          175
83 Leu Lys Gln Gln Val Gly Ser Ser Lys Gln Pro Leu Pro Ile Lys Ala
84          180          185          190
85 Met Leu Arg Ser Leu Pro Ile Trp Ser Ile Cys Leu Gly Cys Phe Ser
86          195          200          205
87 His Gln Trp Leu Val Ser Thr Met Val Val Tyr Ile Pro Thr Tyr Ile
88          210          215          220
89 Ser Ser Val Tyr His Val Asn Ile Arg Asp Asn Gly Leu Leu Ser Ala
90          225          230          235          240
91 Leu Pro Phe Ile Val Ala Trp Val Ile Gly Met Val Gly Gly Tyr Leu
92          245          250          255
93 Ala Asp Phe Leu Leu Thr Lys Lys Phe Arg Leu Ile Thr Val Arg Lys
94          260          265          270
95 Ile Ala Thr Ile Leu Gly Ser Leu Pro Ser Ser Ala Leu Ile Val Ser
96          275          280          285
97 Leu Pro Tyr Leu Asn Ser Gly Tyr Ile Thr Ala Thr Ala Leu Leu Thr
98          290          295          300
99 Leu Ser Cys Gly Leu Ser Thr Leu Cys Gln Ser Gly Ile Tyr Ile Asn
100          305          310          315          320
101 Val Leu Asp Ile Ala Pro Arg Tyr Ser Ser Phe Leu Met Gly Ala Ser
102          325          330          335
103 Arg Gly Phe Ser Ser Ile Ala Pro Val Ile Val Pro Thr Val Ser Gly
104          340          345          350
105 Phe Leu Leu Ser Gln Asp Pro Glu Phe Gly Trp Arg Asn Val Phe Phe
106          355          360          365
107 Leu Leu Phe Ala Val Asn Leu Leu Gly Leu Leu Phe Tyr Leu Ile Phe
108          370          375          380
109 Gly Glu Ala Asp Val Gln Glu Trp Ala Lys Glu Arg Lys Leu Thr Arg
110          385          390          395          400
111 Leu
114 (2) INFORMATION FOR SEQ ID NO: 2:
116 (1) SEQUENCE CHARACTERISTICS:

```

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/991,212

TIME: 10:22:21

Input Set : N:\Crf3\RULE60\09991212.txt

Output Set: N:\CRF3\01152002\I991212.raw

```

117      (A) LENGTH: 1643 base pairs
118      (B) TYPE: nucleic acid
119      (C) STRANDEDNESS: single
120      (D) TOPOLOGY: linear
122      (vii) IMMEDIATE SOURCE:
123          (A) LIBRARY: BRAITUT02
124          (B) CLONE: 754412
126      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
128      AGAACGGTGA GGATGACCGA CGTATAGGCG AGAGCCTAGG TACGCCATGC CAGGTCACCG      60
129      GTCCGGCAAT TCCCGGGTCG ACCCACGCGT CCGCTTGGAG GGACGCTGGG TTCAACTTGA      120
130      AGCCCTTCCA CAGACATTAA GTCGGTGAAA ACCATTCACT AGGAGAGGAG AAACACAATG      180
131      GCCACCAAGA CAGAGTTGAG TCCCACAGCA AGGGAGAGCA AGAACGCACA AGATATGCAA      240
132      GTGGATGAGA CACTGATCCC CAGGAAAGTT CCAAGTTTAT GTTCTGCTCG CTATGGAATA      300
133      GCCCTCGTCT TACATTTCTG CAATTTTACA ACGATAGCAC AAAATGTCAT CATGAACATC      360
134      ACCATGGTAG CCATGGTCAA CAGCACAAGC CCTCAATCCC AGCTCAATGA TTCCTCTGAG      420
135      GTGCTGCCTG TTGACTCATT TGGTGGCCTA AGTAAAGCCC CAAAGAGTCT TCCTGCAAAG      480
136      TCCTCAATAC TTGGGGGTCA GTTTGCAATT TGGGAAAGGT GGGGCCCTCC ACAAGAACGA      540
137      AGCAGACTCT GCAGCATTGC TTTATCAGGA ATGTTACTGG GATGCTTTAC TGCCATCCTC      600
138      ATAGGTGGCT TCATTAGTGA AACCCTTGGG TGGCCCTTTG TCTTCTATAT CTTTGGAGGT      660
139      GTTGGCTGTG TCTGCTGCCT TCTCTGGTTT GTTGTGATTT ATGATGACCC CGTTTCCTAT      720
140      CCATGGATAA GCACCTCAGA AAAAGAATAC ATCATATCCT CCTTGAAACA ACAGGTCGGG      780
141      TCTTCTAAGC AGCCTCTTCC CATCAAAGCT ATGCTCAGAT CTCTACCCAT TTGGTCCATA      840
142      TGTTTAGGCT GTTTCAGCCA TCAATGGTTA GTTAGCACAA TGGTTGTATA CATACCAACT      900
143      TACATCAGCT CTGTGTACCA TGTTAACATC AGAGACAATG GACTTCTATC TGCCCTCCT      960
144      TTTATTGTTG CCTGGGTCAAT AGGCATGGTG GGAGGCTATC TGGCAGATTT CCTTCTAACC      1020
145      AAAAAAGTTA GACTCATCAC TGTGAGGAAA ATTGCCACAA TTTTAGGAAG TCTCCCCTCT      1080
146      TCAGCACTCA TTGTGTCTCT GCCTTACCTC AATTCCGGCT ATATCACAGC AACTGCCTTG      1140
147      CTGACGCTCT CTTGCGGATT AAGCACATTG TGTCAGTCAG GGATTTATAT CAATGTCTTA      1200
148      GATATTGCTC CAAGGTATTC CAGTTTCTCT ATGGGAGCAT CAAGAGGATT TTCGAGCATA      1260
149      GCACCTGTCA TTGTACCCAC TGTCAGCGGA TTTCTTCTTA GTCAGGACCC TGAGTTTGGG      1320
150      TGGAGGAATG TCTTCTTCTT GCTGTTTGCC GTTAACCTGT TAGGACTACT CTTCTACCTC      1380
151      ATATTTGGAG AAGCAGATGT CCAAGAATGG GCTAAAGAGA GAAAACTCAC TCGTTTATGA      1440
152      AGTTATCCCA CCTTGGATGG AAAAGTCATT AGGCACCGTA TTGCATAAAA TAGAAGGCTT      1500
153      CCGTGATGAA AATACCAGTG AAAAGATTTT TTTTTCCTGT GGCTCTTTTC AATTATGAGA      1560
154      TCAGTTCATT ATTTTATTCA GACTTTTTTT TGAGAGAAAT GTAAGATGAA TAAAAATTCA      1620
155      AATAAAATGA TAACTAAGAA TGC      1643
157      (2) INFORMATION FOR SEQ ID NO: 3:
159          (i) SEQUENCE CHARACTERISTICS:
160              (A) LENGTH: 467 amino acids
161              (B) TYPE: amino acid
162              (C) STRANDEDNESS: single
163              (D) TOPOLOGY: linear
165          (vii) IMMEDIATE SOURCE:
166              (A) LIBRARY: GenBank
167              (B) CLONE: 450532
169          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
171      Met Gln Met Asp Asn Arg Leu Pro Pro Lys Lys Val Pro Gly Phe Cys
172      1          5          10          15
173      Ser Phe Arg Tyr Gly Leu Ser Phe Leu Val His Cys Cys Asn Val Ile

```

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/991,212

TIME: 10:22:21

Input Set : N:\Crf3\RULE60\09991212.txt

Output Set: N:\CRF3\01152002\I991212.raw

174				20				25				30				
175	Ile	Thr	Ala	Gln	Arg	Ala	Cys	Leu	Asn	Leu	Thr	Met	Val	Val	Met	Val
176			35					40				45				
177	Asn	Ser	Thr	Asp	Pro	His	Gly	Leu	Pro	Asn	Thr	Ser	Thr	Lys	Lys	Leu
178		50					55					60				
179	Leu	Asp	Asn	Ile	Lys	Asn	Pro	Met	Tyr	Asn	Trp	Ser	Pro	Asp	Ile	Gln
180	65					70					75				80	
181	Gly	Ile	Ile	Leu	Ser	Ser	Thr	Ser	Tyr	Gly	Val	Ile	Ile	Ile	Gln	Val
182					85					90					95	
183	Pro	Val	Gly	Tyr	Phe	Ser	Gly	Ile	Tyr	Ser	Thr	Lys	Lys	Met	Ile	Gly
184			100						105					110		
185	Phe	Ala	Leu	Cys	Leu	Ser	Ser	Val	Leu	Ser	Leu	Leu	Ile	Pro	Pro	Ala
186			115					120					125			
187	Ala	Gly	Ile	Gly	Val	Ala	Trp	Val	Val	Val	Cys	Arg	Ala	Val	Gln	Gly
188		130					135				140					
189	Ala	Ala	Gln	Gly	Ile	Val	Ala	Thr	Ala	Gln	Phe	Glu	Ile	Tyr	Val	Lys
190	145					150					155				160	
191	Trp	Ala	Pro	Pro	Leu	Glu	Arg	Gly	Arg	Leu	Thr	Ser	Met	Ser	Thr	Ser
192					165					170					175	
193	Gly	Phe	Leu	Leu	Gly	Pro	Phe	Ile	Val	Leu	Leu	Val	Thr	Gly	Val	Ile
194				180					185					190		
195	Cys	Glu	Ser	Leu	Gly	Trp	Pro	Met	Val	Phe	Tyr	Ile	Phe	Gly	Ala	Cys
196			195					200					205			
197	Gly	Cys	Ala	Val	Cys	Leu	Leu	Trp	Phe	Val	Leu	Phe	Tyr	Asp	Asp	Pro
198		210					215					220				
199	Lys	Asp	His	Pro	Cys	Ile	Ser	Ile	Ser	Glu	Lys	Glu	Tyr	Ile	Thr	Ser
200	225					230					235				240	
201	Ser	Leu	Val	Gln	Gln	Val	Ser	Ser	Ser	Arg	Gln	Ser	Leu	Pro	Ile	Lys
202					245					250					255	
203	Ala	Ile	Leu	Lys	Ser	Leu	Pro	Val	Trp	Ala	Ile	Ser	Ile	Gly	Ser	Phe
204				260					265					270		
205	Thr	Phe	Phe	Trp	Ser	His	Asn	Ile	Met	Thr	Leu	Tyr	Thr	Pro	Met	Phe
206			275					280					285			
207	Ile	Asn	Ser	Met	Leu	His	Val	Asn	Ile	Lys	Glu	Asn	Gly	Phe	Leu	Ser
208		290					295					300				
209	Ser	Leu	Pro	Tyr	Leu	Phe	Ala	Trp	Ile	Cys	Gly	Asn	Leu	Ala	Gly	Gln
210	305					310					315				320	
211	Leu	Ser	Asp	Phe	Phe	Leu	Thr	Arg	Asn	Ile	Leu	Ser	Val	Ile	Ala	Val
212				325						330					335	
213	Arg	Lys	Leu	Phe	Thr	Ala	Ala	Gly	Phe	Leu	Leu	Pro	Ala	Ile	Phe	Gly
214				340					345					350		
215	Val	Cys	Leu	Pro	Tyr	Leu	Ser	Ser	Thr	Phe	Tyr	Ser	Ile	Val	Ile	Phe
216			355					360					365			
217	Leu	Ile	Leu	Ala	Gly	Ala	Thr	Gly	Ser	Phe	Cys	Leu	Gly	Gly	Val	Phe
218		370					375					380				
219	Ile	Asn	Gly	Leu	Asp	Ile	Ala	Pro	Arg	Tyr	Phe	Gly	Phe	Ile	Lys	Ala
220	385					390					395				400	
221	Cys	Ser	Thr	Leu	Thr	Gly	Met	Ile	Gly	Gly	Leu	Ile	Ala	Ser	Thr	Leu
222					405					410					415	

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/991,212

TIME: 10:22:21

Input Set : N:\Crf3\RULE60\09991212.txt

Output Set: N:\CRF3\01152002\I991212.raw

```

223 Thr Gly Leu Ile Leu Lys Gln Asp Pro Glu Ser Ala Trp Phe Lys Thr
224           420           425           430
225 Phe Ile Leu Met Ala Ala Ile Asn Val Thr Gly Leu Ile Phe Tyr Leu
226           435           440           445
227 Ile Val Ala Thr Ala Glu Ile Gln Asp Trp Ala Lys Glu Lys Gln His
228           450           455           460
229 Thr Arg Leu
230 465
232 (2) INFORMATION FOR SEQ ID NO: 4:
233 (i) SEQUENCE CHARACTERISTICS:
234 (A) LENGTH: 560 amino acids
235 (B) TYPE: amino acid
236 (C) STRANDEDNESS: single
237 (D) TOPOLOGY: linear
240 (vii) IMMEDIATE SOURCE:
241 (A) LIBRARY: GenBank
242 (B) CLONE: 507415
244 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
246 Met Glu Phe Arg Gln Glu Glu Phe Arg Lys Leu Ala Gly Arg Ala Leu
247 1           5           10           15
248 Gly Arg Leu His Arg Leu Leu Glu Lys Arg Gln Glu Gly Ala Glu Thr
249           20           25           30
250 Leu Glu Leu Ser Ala Asp Gly Arg Pro Val Thr Thr His Thr Arg Asp
251           35           40           45
252 Pro Pro Val Val Asp Cys Thr Cys Phe Gly Leu Pro Arg Arg Tyr Ile
253           50           55           60
254 Ile Ala Ile Met Ser Gly Leu Gly Phe Cys Ile Ser Phe Gly Ile Arg
255           65           70           75           80
256 Cys Asn Leu Gly Val Ala Ile Val Ser Met Val Asn Asn Ser Thr Thr
257           85           90           95
258 His Arg Gly Gly His Val Val Val Gln Lys Ala Gln Phe Asn Trp Asp
259           100          105          110
260 Pro Glu Thr Val Gly Leu Ile His Gly Ser Phe Phe Trp Gly Tyr Ile
261           115          120          125
262 Val Thr Gln Ile Pro Gly Gly Phe Ile Cys Gln Lys Phe Ala Ala Asn
263           130          135          140
264 Arg Val Phe Gly Phe Ala Ile Val Ala Thr Ser Thr Leu Asn Met Leu
265           145          150          155          160
266 Ile Pro Ser Ala Ala Arg Val His Tyr Gly Cys Val Ile Phe Val Arg
267           165          170          175
268 Ile Leu Gln Gly Leu Val Glu Gly Val Thr Tyr Pro Ala Cys His Gly
269           180          185          190
270 Ile Trp Ser Lys Trp Ala Pro Pro Leu Glu Arg Ser Arg Leu Ala Thr
271           195          200          205
272 Thr Ala Phe Cys Gly Ser Tyr Ala Gly Ala Val Val Ala Met Pro Leu
273           210          215          220
274 Ala Gly Val Leu Val Gln Tyr Ser Gly Trp Ser Ser Val Phe Tyr Val
275           225          230          235          240
276 Tyr Gly Ser Phe Gly Ile Phe Trp Tyr Leu Phe Trp Leu Leu Val Ser

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/991,212

DATE: 01/15/2002

TIME: 10:22:22

Input Set : N:\Crf3\RULE60\09991212.txt

Output Set: N:\CRF3\01152002\I991212.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]